Entropic Independence: Optimal mixing of down-up walk

Thuy Duong "June" Vuong

UChicago seminar

Joint work with Nima Anari, Vishesh Jain, Frederic Koehler, Huy T. Pham




Sampling from a distribution

- Given access to density function $\mu: \Omega \to \mathbb{R}_{\geq 0}$, output x in Ω s.t.

$$\mathbb{P}[x] \propto \mu(x)$$

E.g.: (sampling problems)

- random spanning tree [Aldous-Broder'90, Colbourn-Myrvold-Neufeld'96, Kelner-Madry'09, Madry-Straszak-Tarnawski'15, Schild'18, Anari-Liu-OveisGharan-Vinzant-**V.**—STOC'21]
- matroid bases [Anari-Liu-OveisGharan-Vinzant—STOC'19,Cryan-Guo-Mousa—FOCS'19]
- perfect matching? For bipartite graph: [Jerrum-Sinclair-Vigoda'04]

Sampling from a distribution

- Given access to density function $\mu: \Omega \to \mathbb{R}_{\geq 0}$, output x in Ω s.t.
$$\mathbb{P}[x] \propto \mu(x)$$
- Sufficient to approximately sample i.e. output x according to $\widehat{\mathbb{P}}$ s.t.
$$d_{TV}\left(\mathbb{P}, \widehat{\mathbb{P}}\right) = \sum |\widehat{\mathbb{P}}(x) - \mathbb{P}(\mathrm{x})| < 0.01$$

Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

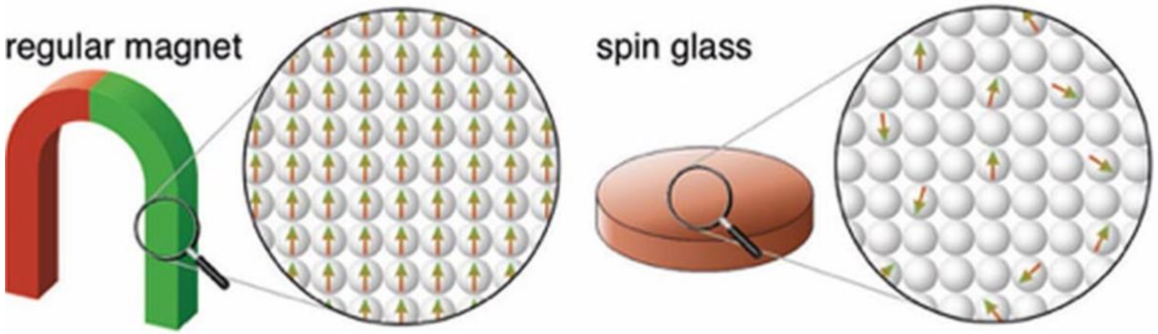
Ising models

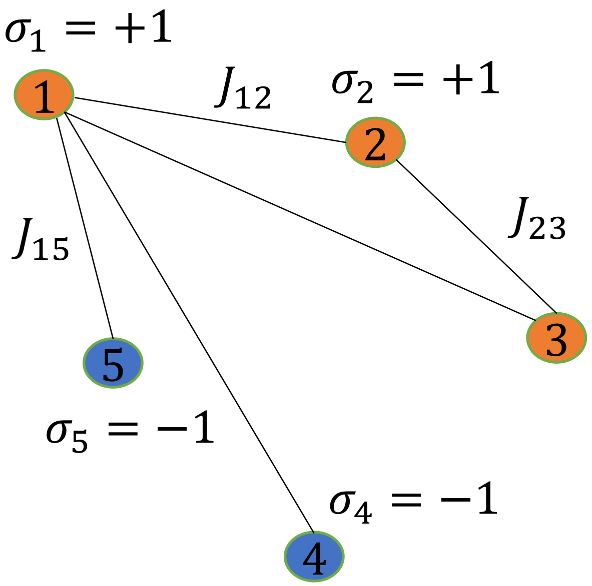
Counting/Sampling:

Structure of materials

Neural Networks--Hopfield Model

Interacting particle process (Ligett)





Interaction matrix J

$$\begin{pmatrix} & J_{12} & \\ J_{12} & & \\ & & \end{pmatrix}$$

External field h

$$\begin{pmatrix} h_1 & h_2 & \dots & h_5 \end{pmatrix}$$

$$\mu_{J,h}(\sigma) = \exp\left(\sum_{i<j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i\right)$$

Hardcore model

Counting/sampling independent sets of graph G = G (V,E) with max degree Δ

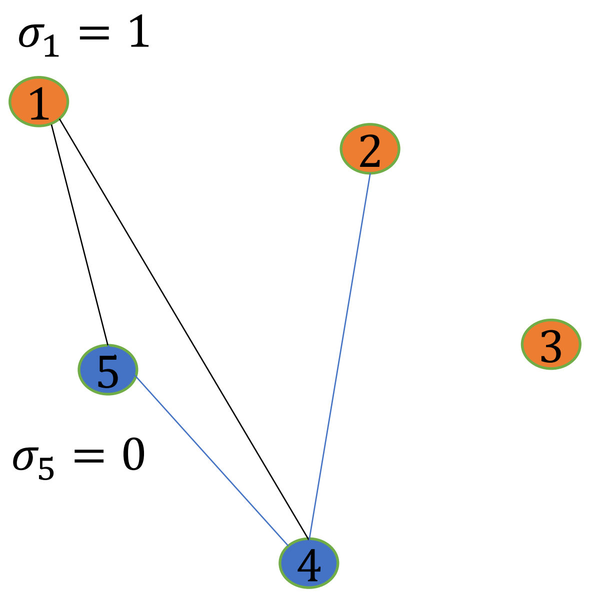
MIS is NP-hard

$\lambda \geq \lambda_\Delta \approx \frac{e}{\Delta}$: NP-hard to count/sample

$\lambda < \lambda_\Delta(1-\delta)$ (tree-unique region):

$\mu_{G,\lambda}$ has correlation decay. Easy to sample?

Many recent results



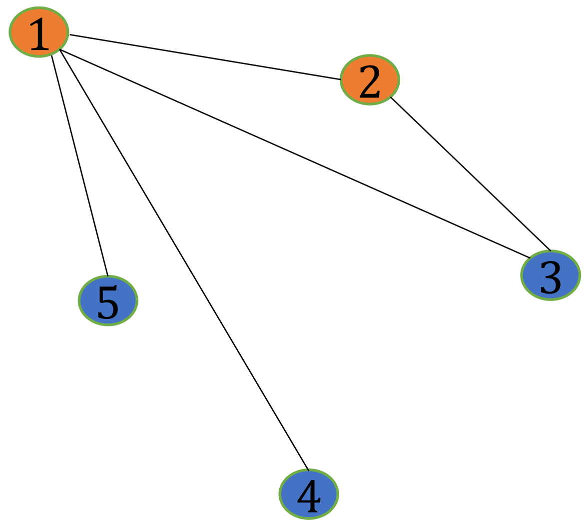
$$\mu_{G,\lambda}(\sigma) = \prod_{(i,j)\in E(G)} 1[\sigma_i\sigma_j = 0] \prod_{i:\sigma_i=1} \lambda_i$$

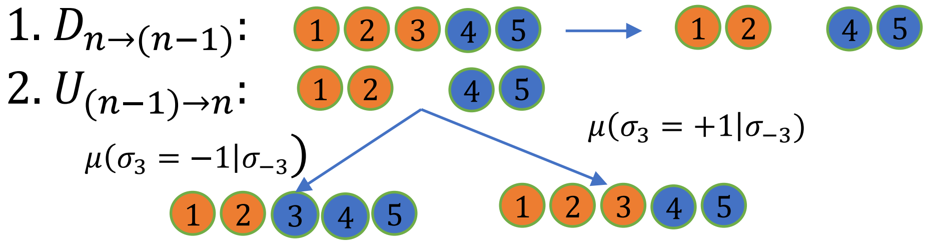
Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

Sampling using Glauber dynamics

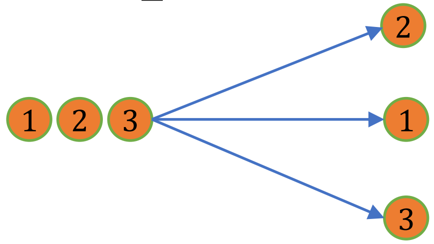
- Start at distribution μ_0, apply transition rule for T steps to reach desired distribution μ.
- Want: $T = O(n \log n)$ i.e. optimal mixing time.
- Need: theory to bound mixing time.



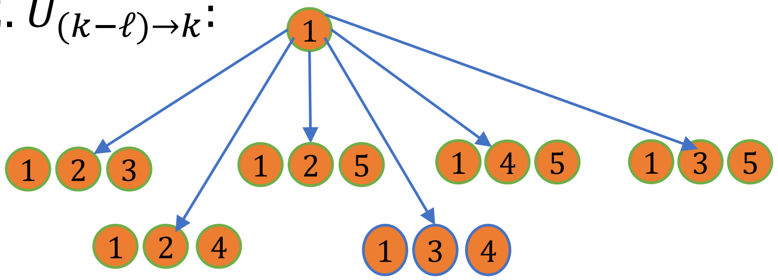


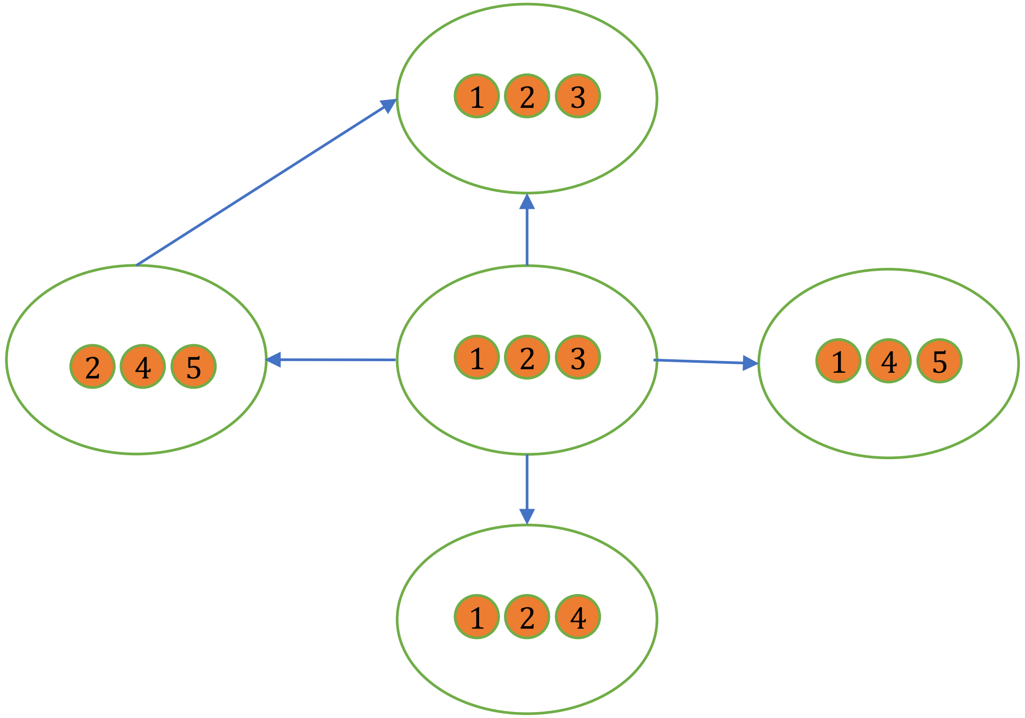
Multi-steps down-up walk for $\mu: \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$

1. $D_{k\to(k-\ell)}$:



2. $U_{(k-\ell)\to k}$:





$$P = D_{k\to(k-\ell)} U_{(k-\ell)\to k}$$

Why study (multi-step) down-up walks?

- 1-step down-up walks ≡ Glauber dynamics, basis exchange walks to sample matroid bases
- 2-step down-up walks: sample matchings in planar graph

[Alimohammadi-Anari-Shiragur-**V.**—STOC'21]

- Block Glauber dynamics [Chen-Liu-Vigoda—STOC'21]
- Field dynamics (to sample from hardcore models) [Chen-Yin-Feng-Zhang—FOCS'21]

Overview

1. Motivation
    - Ising and hardcore model
    - Glauber dynamics
    - Multi-step down-up walks
    - Markov chain and mixing time
2. Entropic Independence
    - Definition
    - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
    - Local-to-global argument
    - Glauber dynamics for Ising/hardcore models

Markov chain and mixing time

- Markov chain with transition matrix P with stationary dist. μ

$$\nu \to \nu P \to \nu P^2 \to \cdots \to \mu = \mu P$$

- Distance between probability distribution (f-divergence)

$$\mathcal{D}_f(\nu||\mu) = \mathbb{E}_\mu\left[f\left(\frac{\nu(x)}{\mu(x)}\right)\right] - f\left(\mathbb{E}_\mu\left[\frac{\nu(x)}{\mu(x)}\right]\right) \geq 0$$

- To bound number of steps till convergence (T_{mix}), need to show P contract D_f

$$\mathcal{D}_f(\nu P||\mu P) \leq \left(1 - \rho_f\right)\mathcal{D}_f(\nu||\mu)$$

f-divergence contraction vs. mixing time

Variance contraction ($f = x^2$)

- $T_{mix} \le \rho_{\chi^2}^{-1} \log \min \mu(x)^{-1}$
- $\rho_{\chi^2} = 1 - \lambda_2(P)$

f-divergence contraction vs. mixing time

Variance contraction ($f = x^2$)

- $T_{mix} \leq \rho_{\chi^2}^{-1} \log \min \mu(x)^{-1}$
- $\rho_{\chi^2} = 1 - \lambda_2(P)$

Entropy contraction ($f = x \log x$)

- $T_{mix} \leq \rho_{KL}^{-1} \log \log \ \min \mu(x)^{-1}$
- $\mathcal{D}_{x \log x} = \mathcal{D}_{KL}$

f-divergence contraction vs. mixing time

Variance contraction ($\boldsymbol{f = x^2}$)

- $T_{mix} \leq \rho_{\chi^2}^{-1} \log \min \mu(x)^{-1}$
- $\rho_{\chi^2} = 1 - \lambda_2(P)$

Entropy contraction ($\boldsymbol{f = x\, log\, x}$)

- $T_{mix} \leq \rho_{KL}^{-1} \log \log \ \min \mu(x)^{-1}$
- $\mathcal{D}_{x \log x} = \mathcal{D}_{KL}$

Typically for Glauber dynamics: $\rho_{\chi^2} = \rho_{KL} = \frac{1}{n}$

but $\log \min \mu(x)^{-1} \approx n$. Bounding $\rho_{KL} \Rightarrow$ quadratic improvement on T_{mix}

Bonus: $\rho_{KL} = 1/n \Rightarrow$ concentration of Lipschitz functions

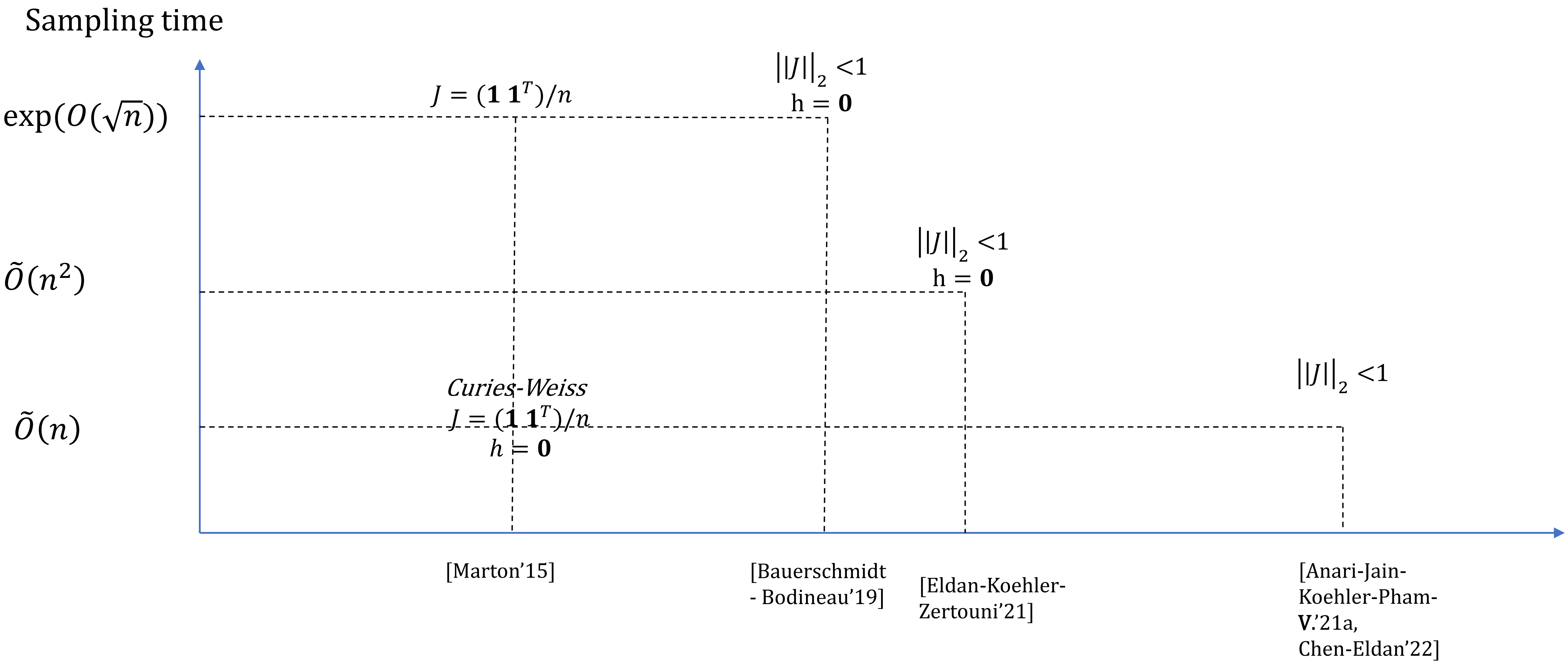
It is hard to bound ρ_{KL}!

Multi-steps down-up walks

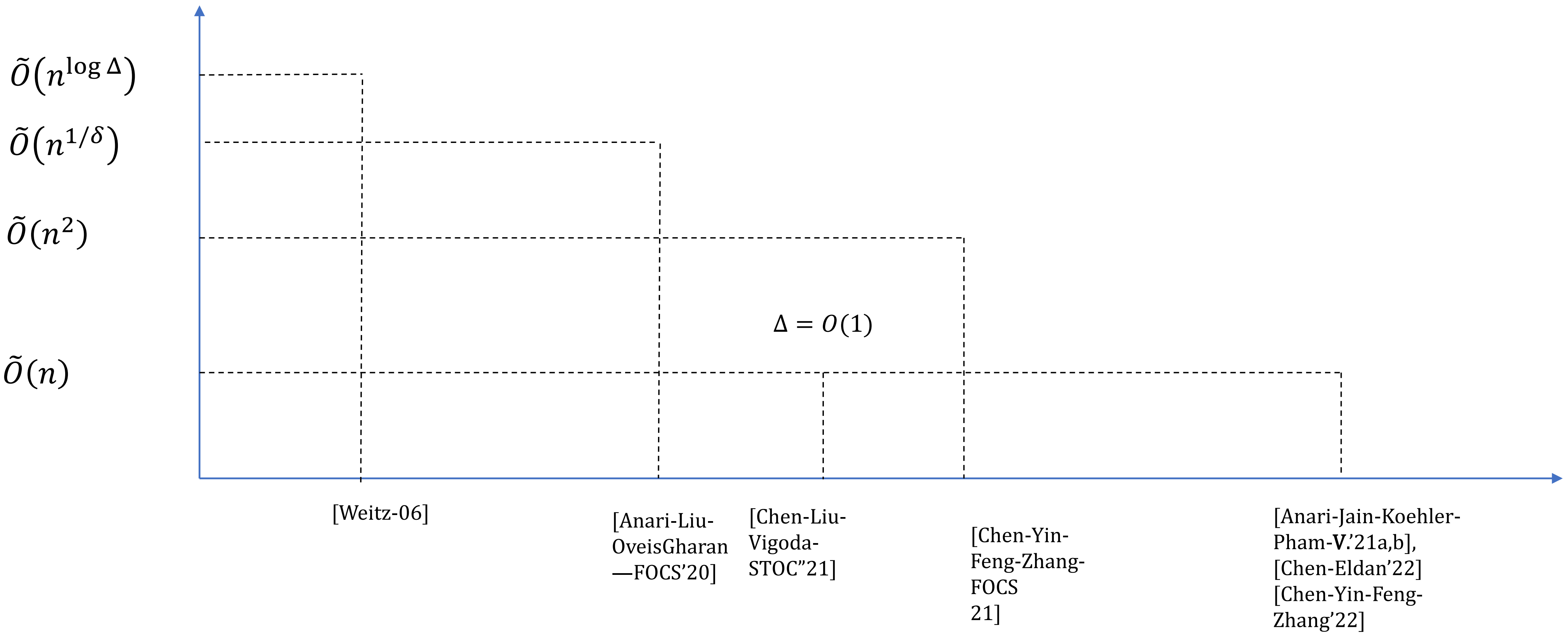
- Transition matrix $P = D_{k\to(k-\ell)}U_{(k-\ell)\to k}$
- Reversible
- Converge to μ (μ is used to define up-operator)
- Mixing time is controlled by entropy contraction of $D_{k\to(k-\ell)}$

$$\mathcal{D}_{KL}(\nu D_{k\to(k-\ell)}||\mu D_{k\to(k-\ell)}) \leq (1-\rho)\mathcal{D}_f(\nu||\mu)$$

Sampling from Ising models



Sampling from hardcore model $(\lambda < \lambda_\Delta(1-\delta))$



Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Multi-step down-up walks
   - Local-to-global argument
   - Glauber dynamics for Ising models

Spectral Independence

$D_{k\to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-spectral independence ⇔ ∀ν: $\mathcal{D}_{\chi^2}(\nu || \mu) \geq \alpha k \mathcal{D}_{\chi^2}(\nu D_{k\to 1} || \mu D_{k\to 1})$ (1)

$$(1) \Leftrightarrow \lambda_2(U_{1\to k} D_{k\to 1}) = \lambda_2(U_{1\to 2} D_{2\to 1}) \leq 1 - \alpha/k \Leftrightarrow ||\Psi_\mu^{corr}||_2 \leq \frac{1}{\alpha}$$

Spectral Independence

$D_{k\to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-spectral independence ⇔ $\forall \nu$: $\mathcal{D}_{\chi^2}(\nu||\mu) \geq \alpha k \mathcal{D}_{\chi^2}(\nu D_{k\to 1}||\mu D_{k\to 1})$ (1)

Spectral independence ⇒ contraction of $\mathcal{D}_{\chi^2}$ by $D_{k\to(k-\ell)}$

? ⇒ contraction of $\mathcal{D}_{KL}$ by $D_{k\to(k-\ell)}$

Entropic Independence

$D_{k\to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-spectral independence ⇔ $\forall \nu$: $\mathcal{D}_{\chi^2}(\nu || \mu) \geq \alpha k \mathcal{D}_{\chi^2}(\nu D_{k\to 1} || \mu D_{k\to 1})$

$\frac{1}{\alpha}$-entropic independence ⇔ $\forall \nu$: $\mathcal{D}_{KL}(\nu || \mu) \geq \alpha k \mathcal{D}_{KL}(\nu D_{k\to 1} || \mu D_{k\to 1})$

Entropic Independence

$D_{k \to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-spectral independence ⇔ $\forall \nu$: $\mathcal{D}_{\chi^2}(\nu || \mu) \geq \alpha k \mathcal{D}_{\chi^2}(\nu D_{k \to 1} || \mu D_{k \to 1})$

$\frac{1}{\alpha}$-entropic independence ⇔ $\forall \nu$: $\mathcal{D}_{KL}(\nu || \mu) \geq \alpha k \mathcal{D}_{KL}(\nu D_{k \to 1} || \mu D_{k \to 1})$

Scaling of μ by λ:

$$\lambda * \mu(S) = \mu(S) \prod_{i \in S} \lambda_i$$

Main theorem:

$\frac{1}{\alpha}$-spectral independence of $\lambda * \mu \forall$ (α-FLC) ⇒ $\frac{1}{\alpha}$-entropic independence of μ

Main theorem

$\frac{1}{\alpha}$-spectral independence of $\lambda * \mu \forall \lambda \in \mathbb{R}^n_{\geq 0}$

$\Rightarrow \frac{1}{\alpha}$-entropic independence of μ

Entropic independence

$D_{k\to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-entropic independence ⇔ ∀ν: $\mathcal{D}_{KL}(\nu || \mu) \geq \alpha k \mathcal{D}_{KL}(\nu D_{k\to 1} || \mu D_{k\to 1})$

Why $D_{k\to 1}$ instead of $D_{2\to 1}$?

- $D_{2\to 1}$ has no entropy contraction for natural distributions of interest
- $D_{2\to 1}$ has contraction only for restricted case: distribution on O(1)-bounded degree graphs with bounded marginals [Chen-Liu-Vigoda—STOC'21], uniform distribution over matroid bases [Cryan-Guo-Mousa—FOCS'19]

From EI to optimal mixing time

$D_{k\to 1}(S)$: sample $i \in S$ uniformly

$\frac{1}{\alpha}$-entropic independence ⇔ ∀ν: $\mathcal{D}_{KL}(\nu||\mu) \geq \alpha k \mathcal{D}_{KL}(\nu D_{k\to 1}||\mu D_{k\to 1})$

Local-to-global (similar to [Alev-Lau—STOC'21]) ⇒ optimal mixing for Glauber dynamics on Ising/hardcore model and other local walks

Main theorem

$\frac{1}{\alpha}$-spectral independence of $\lambda * \mu \forall$ (α-Fractionally Log Concave)
$\Rightarrow \frac{1}{\alpha}$-entropic independence of μ

Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

Generating polynomial

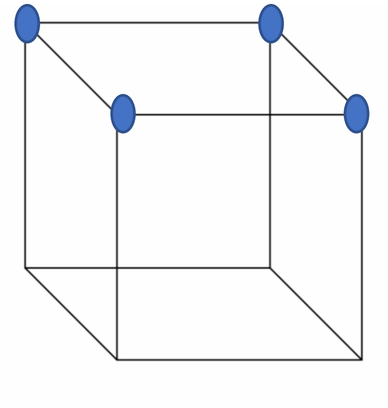
Generating polynomial of $\mu: \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$

$$f_\mu(z_1, \ldots, z_n) := \sum \mu(S) \prod_{i \in S} z_i$$

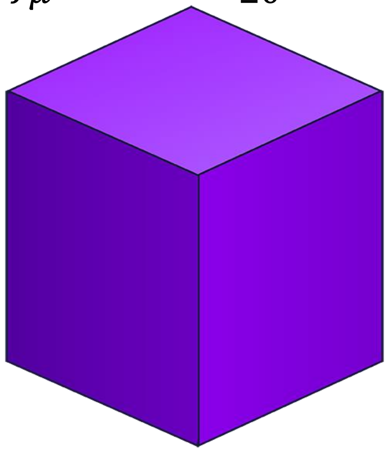
Scaling of μ by external field $\lambda \in \mathbb{R}^n_{\geq 0}$:

$$\lambda * \mu(S) = \mu(S) \prod_{i \in S} \lambda_i$$

$\mu: \{0,1\}^n \to \mathbb{R}_{\geq 0}$



$f_\mu: \mathbb{R}^n \to \mathbb{R}_{\geq 0}$



EI/FLC vs. geometry of polynomial

$\frac{1}{\alpha}$-entropic independence: $f_\mu(z_1^\alpha, \ldots, z_n^\alpha)^{1/k\alpha} \leq \sum z_i p_i^\mu$ for $z_i \in (0, +\infty)$ (2)

EI/FLC vs. geometry of polynomial

$\frac{1}{\alpha}$-entropic independence: $f_\mu(z_1^\alpha, \dots, z_n^\alpha)^{1/k\alpha} \le \sum z_i p_i^\mu$ for $z_i \in (0, +\infty)$ (2)

$\frac{1}{\alpha}$-spectral independence: $f_\mu(z_1^\alpha, \dots, z_n^\alpha)^{1/k\alpha} \le \sum z_i p_i^\mu$ for $z_i \in (1-\epsilon, 1+\epsilon)$

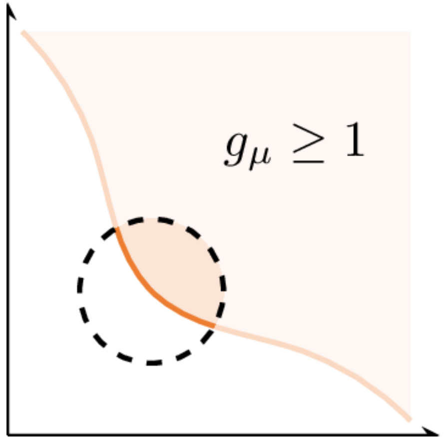
α-fractional-log concave: $f_{\lambda * \mu}(z_1^\alpha, \dots, z_n^\alpha)^{1/k\alpha} \le \sum z_i p_i^{\lambda * \mu}$ for $\lambda_i, z_i \in (0, +\infty)$

α-fractional-log concave $\Rightarrow$ (2) $\underset{(*)}{\Leftrightarrow}$ $\frac{1}{\alpha}$-entropic independence

Geometry of Polynomials

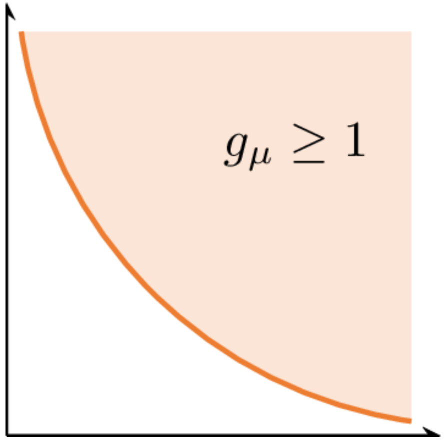
$$\mathrm{h} = \log \mathrm{f}(z_1^\alpha, \dots, z_n^\alpha), \mathrm{p_i} = \mathbb{P}_{S\sim\mu}[i \in S] = \partial_i\, h(1, \dots, 1)$$

$$\mathrm{h\ concave} \Leftrightarrow g := f(z_1^\alpha, \dots, z_n^\alpha)^{1/k\alpha}\ \mathrm{concave}$$

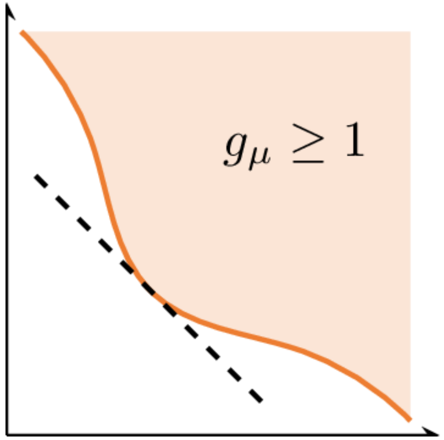


$\frac{1}{\alpha}$-Spectral independence

g concave
for $z_i = 1$



α-Fractional-log-concave
g concave in $\mathbb{R}_+^n$



$\frac{1}{\alpha}$-Entropic independence

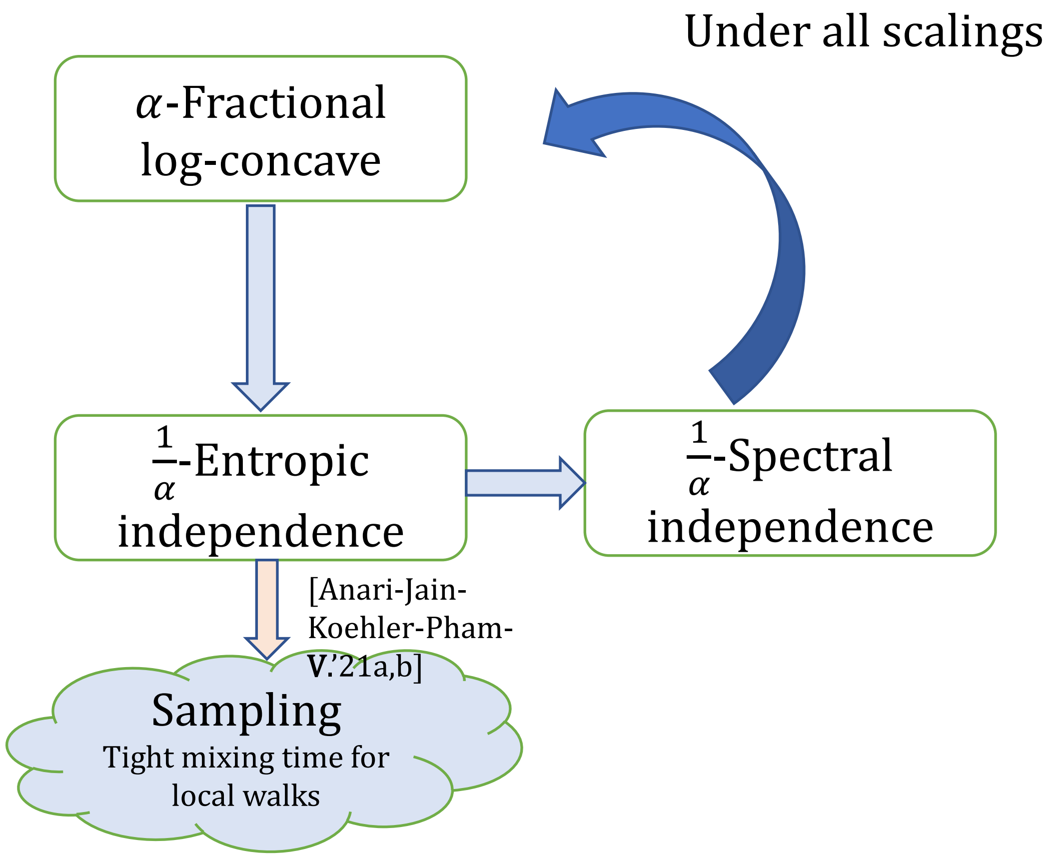
$f(z_1^\alpha, \dots, z_n^\alpha)^{1/k\alpha} \leq \sum z_i p_i$
for $z_i \in (0, C)$

[Anari-Liu-OveisGharan --FOCS'20]

[Alimohammadi-Anari-Shiragur-**V.**--STOC'21]

[Anari-Jain-Koehler-Pham-**V.**'21a,b]

Geometry of polynomials



Proof of (*)

- Minimize $\mathcal{D}_{KL}(\nu||\mu) = \sum \nu(S) \log \nu(S)/\mu(S)$ s.t. $\nu D_{k\to 1}$=q
- Minimizer: $\nu(S) = \mu(S)\lambda^S = \mu(S) \prod_{i\in S} z\_i$
- Lagrange multiplier:

$$\min \{\mathcal{D}_{KL}(\nu||\mu) | \nu D_{\mathrm{k}\to 1} = q\} \geq \min \Phi(\nu) := \sum \nu(S) \log \frac{\nu(S)}{\mu(S)} - \lambda(\textstyle\sum_S \nu(S) - 1) - \sum_i \lambda_i (\sum_{S \ni i} \nu(S) - kq_i)$$

By duality in convex programming:

$$\min \Phi(\nu) = \min(-\ln f_\mu(z_1, \dots, z_n) + \textstyle\sum_i kq_i \log z_i)$$

Proof of (*)

$$\exp(-\frac{\Phi(\nu)}{k\alpha}) \leq \frac{\sum p_i y_i}{y_1^{q_1} \dots y_n^{q_n}}$$

$$y_i = \frac{q_i}{p_i} \Rightarrow \exp(-\min\frac{\Phi(\nu)}{k\alpha}) \leq \frac{\sum q_i}{\left(\frac{q_1}{p_1}\right)^{q_1} \dots \left(\frac{q_n}{p_n}\right)^{q_n}}$$

$$\frac{\min\{\mathcal{D}_{KL}(\nu||\mu) | \nu D_{\mathrm{k}\to 1} = q\}}{k\alpha} \geq \min\frac{\Phi(\nu)}{k\alpha} \geq \sum q_i \log\frac{q_i}{p_i} = \mathcal{D}_{KL}(\nu D_{k\to 1}||\mu D_{k\to 1})$$

Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

From EI to optimal mixing time

- If $\mu: \binom{[n]}{k} \to \mathbb{R}_{\geq 0}$ and its conditionals are $\frac{1}{\alpha}$-EI then for $\ell = \lceil \frac{1}{\alpha} \rceil$

$$\left(1 - \frac{1}{k^{\frac{1}{\alpha}}}\right)\mathcal{D}_{KL}(\nu||\mu) \geq \mathcal{D}_{KL}\left(\nu D_{k\to(k-\ell)}\middle|\middle|\mu D_{k\to(k-\ell)}\right)$$

Thus $\lceil \frac{1}{\alpha} \rceil$-steps down-up walk has mixing time $\tilde{O}(k^{\frac{1}{\alpha}})$

From EI to optimal mixing time

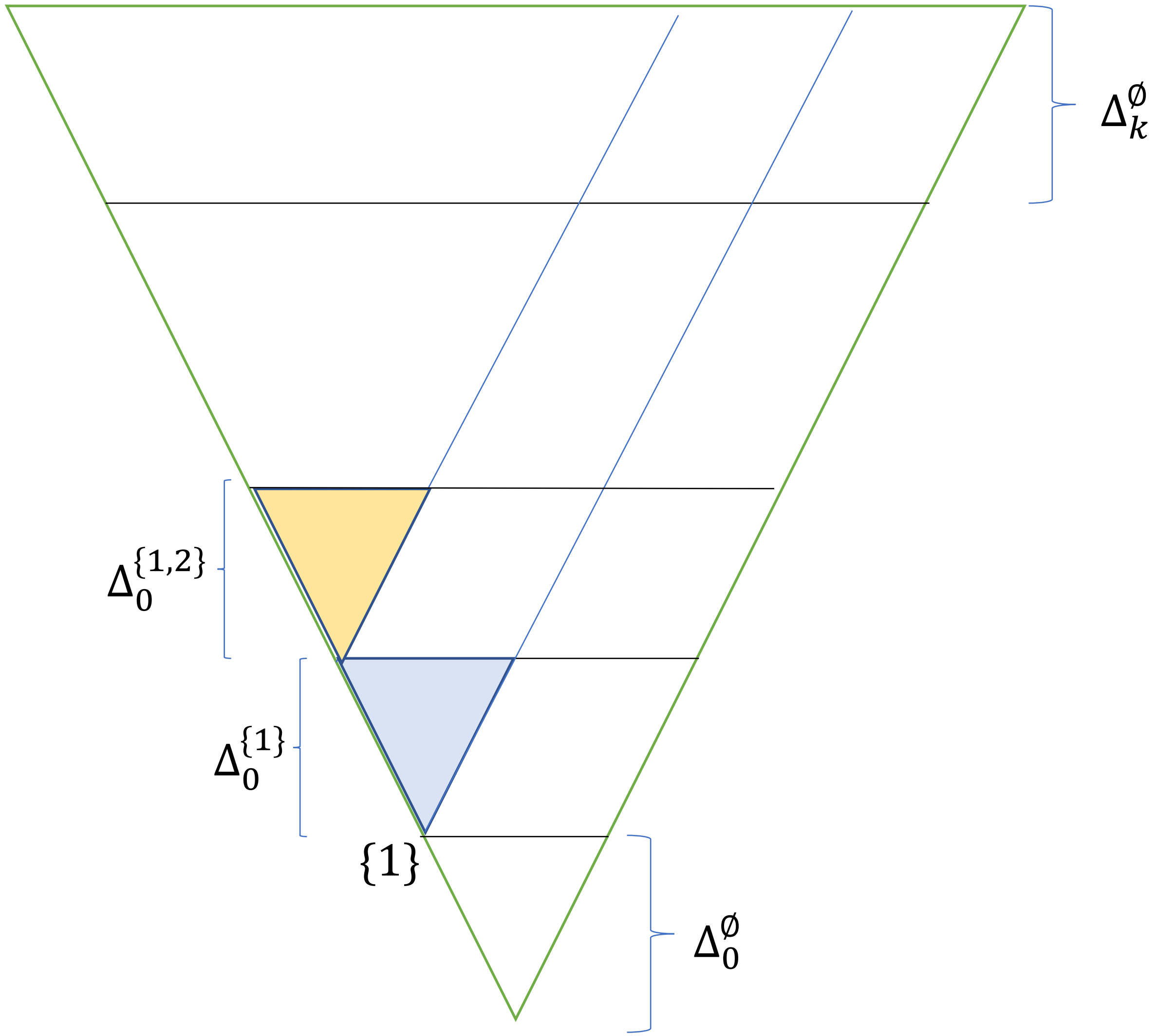
- If μ is α-FLC then for $\ell = \lceil \frac{1}{\alpha} \rceil$

$$\left(1 - \frac{1}{k^{\frac{1}{\alpha}}}\right)\mathcal{D}_{KL}(\nu||\mu) \geq \mathcal{D}_{KL}\left(\nu D_{k\to(k-\ell)}\middle|\middle|\mu D_{k\to(k-\ell)}\right)$$

Thus $\lceil \frac{1}{\alpha} \rceil$-steps down-up walk has mixing time $\tilde{O}(k^{\frac{1}{\alpha}})$

Extend [Cryan-Guo-Mousa—STOC'20] for $\alpha < 1$.

Local to global (Similar to [Alev-Lau—STOC'21])



$$\underbrace{\mathcal{D}_{KL}(\nu||\mu)}_{\sum_{i=0}^{k}\Delta_i^{\emptyset}} \geq \alpha k \underbrace{\mathcal{D}_{KL}(\nu D_{k\to1}||\mu D_{k\to1})}_{\Delta_0^{\emptyset}}$$

$$\sum_{i=0}^{k-1}\Delta_i^{\{j\}} \geq \alpha(k-1)\,\Delta_0^{\{j\}}$$

$$\underbrace{\sum_j \sum_{i=0}^{k-1}\Delta_i^{\{j\}}}_{\sum_{i=1}^{k}\Delta_i^{\emptyset}} \geq \alpha(k-1)\underbrace{\sum_j \Delta_0^{\{j\}}}_{\Delta_1^{\emptyset}}$$

Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

Glauber dynamics on Ising models

- [Eldan-Koehler-Zertouni'21]: Reduce to interaction matrix J of rank 1 i.e. $J = u^T u$
- Naïve local-to-global ⇒ $O\left(n^{1/(1-||u||_2^2)}\right)$-mixing of $1/\left(1-||u||_2^2\right)$-steps down up walk
- Need: $O\left(n/(1-||u||_2^2)\right)$-mixing of Glauber dynamics (1-step down-up walk)

Induction: $\left(1-\frac{1-||u||_2^2}{n}\right)\mathcal{D}_{KL}(\nu||\mu) \geq \mathcal{D}_{KL}(\nu D||\mu D)$

with $D = D_{n\to(n-1)}$

- $\mu(.\,|\sigma_i = +1)$ is Ising model with $J = u_{-i}^T u_{-i}$
- Apply induction hypothesis to $\mu^{(+i)} = \mu(.\,|\sigma_i = +1)$ gives

$$\left(1-\frac{1-||u_{-i}||_2^2}{n-1}\right)\mathcal{D}_{KL}(\nu^{(+i)}||\mu^{(+i)}) \geq \mathcal{D}_{KL}(\nu^{(+i)}D||\mu^{(+i)}D)$$

Induction: $\left(1-\frac{1-||u||_2^2}{n}\right)\mathcal{D}_{KL}(\nu||\mu) \geq \mathcal{D}_{KL}(\nu D||\mu D)$ with $D = D_{n\to(n-1)}$

- $\mu(.\,|\sigma_i = +1)$ is Ising model with $J = u_{-i}^T u_{-i}$
- Apply induction hypothesis to $\mu^{(+i)} = \mu(.\,|\sigma_i = +1)$ gives

$$\left(1-\frac{1-||u_{-i}||_2^2}{n-1}\right)\mathcal{D}_{KL}(\nu^{(+i)}||\mu^{(+i)}) \geq \mathcal{D}_{KL}(\nu^{(+i)}D||\mu^{(+i)}D)$$

- $\nu_i(+1)\mathcal{D}_{KL}(\nu^{(+i)}||\mu^{(+i)}) + \nu_i(-1)\mathcal{D}_{KL}(\nu^{(-i)}||\mu^{(-i)}) = \mathcal{D}_{KL}(\nu||\mu) - \mathcal{D}_{KL}(\nu_i||\mu_i)$
- $\left(1-\frac{1-||u_{-i}||_2^2}{n-1}\right)\mathcal{D}_{KL}(\nu||\mu) + \frac{1-||u_{-i}||_2^2}{n-1}\mathcal{D}_{KL}(\nu_i||\mu_i) \geq \mathcal{D}_{KL}(\nu D||\mu D)$

Where for $\mu,\nu:\{\pm1\}^n \to \mathbb{R}_{\geq 0}$, $\mu_i,\nu_\mathrm{i}:\{\pm1\} \to \mathbb{R}_{\geq 0}$ are marginal distributions of i-th coordinate.

Induction (continue)

- Average over i gives

$$\mathcal{D}_{KL}(\nu D||\mu D) \le \frac{1}{n}\left[\sum_i \left(1 - \frac{1-\left||u_{-i}|\right|_2^2}{n-1}\right)\mathcal{D}_{KL}(\nu||\mu) + \frac{1-\left||u_{-i}|\right|_2^2}{n-1}\mathcal{D}_{KL}(\nu_i||\mu_i)\right]$$

EI

$$\le \left(1 - \frac{1}{n-1} + \frac{\left||u|\right|_2^2}{n}\right)\mathcal{D}_{KL}(\nu||\mu) \qquad + \frac{1}{n(n-1)}\mathcal{D}_{KL}(\nu||\mu)$$

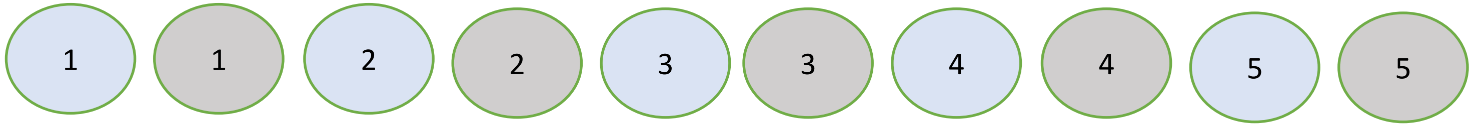
Non-uniform entropic independence

For $\mu, \nu: \{\pm 1\}^n \to \mathbb{R}_{\geq 0}$ let $\mu_i, \nu_{\mathrm{i}}: \{\pm 1\} \to \mathbb{R}_{\geq 0}$ be marginal distribution of i-th coordinate. $\mu \equiv \mu^{hom}: \binom{[n] \cup [\bar{n}]}{n} \to \mathbb{R}_{\geq 0}$

We say μ is $(\alpha_1, \dots, \alpha_n)$-entropic independence (EI) if

$$\forall \mathrm{v}: \ \mathcal{D}_{KL}(\nu || \mu) \geq \sum_i \alpha_i \, \mathcal{D}_{KL}(\nu_i || \mu_i)$$

$(\alpha_1, \dots, \alpha_n)$-FLC $\Rightarrow$ $(\alpha_1, \dots, \alpha_n)$-EI



$$\mu_1(+1) + \mu_1\,(-1) = 1$$

Entropic independence of Ising models

- For $\mu \equiv$ Ising model with $J = u^T u$, show $(\alpha_1, \dots, \alpha_n)$-FLC/EI with $\alpha_i = \left(1 - \left|\left|u_{-i}\right|\right|_2^2\right) = 1 - \sum_{j \neq i} u_j^2$

Entropic independence of Ising models

- For $\mu \equiv$ Ising model with $J = u^T u$, show $(\alpha_1, \ldots, \alpha_n)$-FLC/EI with $\alpha_i = \left(1 - \left||u_{-i}|\right|_2^2\right) = 1 - \sum_{j \neq i} u_j^2$
- Dobrushin matrix: $\sigma^i \equiv \sigma$ with i-th coordinate flipped

$$R_{ij} = \max_{\sigma_{-i}} d_{TV}(\mu(.|\sigma_{-i}), \mu\left(.\middle|\sigma_{-i}^{j}\right)) \leq |u_i u_j|$$

Entropic independence of Ising models

- For $\mu \equiv$ Ising model with $J = u^T u$, show $(\alpha_1, \ldots, \alpha_n)$-FLC/EI with $\alpha_i = \left(1 - \left||u_{-i}|\right|_2^2\right)$
- Dobrushin matrix: $\sigma^i \equiv \sigma$ with i-th coordinate flipped

$$R_{ij} = \max_{\sigma_{-i}} d_{TV}(\mu(.\,|\sigma_{-i}), \mu\left(.\left|\sigma_{-i}^{j}\right.\right)) \leq |u_i u_j|$$

- $U := diag(|u_1|, \ldots, |u_n|)$ then $||URU^{-1}||_1 \leq ||u||_2^2$

+[Blanca-Caputo-Chen-Parisi-Stefankovic-Vigoda’21,Liu—RANDOM’21]⇒ $(1 - \left||u|\right|_2^2)$-FLC

+[AJKPV’21a]: $(\alpha_1, \ldots, \alpha_n)$-FLC

(continue)

- Influence matrix $\Psi_\mu^{\inf}(i,j) = \mu\left(\sigma_j = +1 \middle| \sigma_i = +1\right) - \mu\left(\sigma_j = +1 \middle| \sigma_i = +1\right)$
- Let $\mu^{(\pm i)} = \mu(.\,|\sigma_i = \pm 1)$ then
- $\sum \left| u_j \Psi_\mu^{\inf}(i,j) \right| \leq \frac{n-1}{\alpha_i} \max_{\sigma_{-i}} \sum_j |\mathrm{u_j}| \left(\mathrm{P}_{\mu^{(+i)}}\left(\sigma_{-i} \to \sigma_{-i}^{j}\right) - \mathrm{P}_{\mu^{(-i)}}\left(\sigma_{-i} \to \sigma_{-i}^{j}\right) \right) |$

$$\leq \frac{1}{\alpha_i} \sum_j |\mathrm{u_j}| R_{ij} \leq \frac{1}{\alpha_i} u_i \left| |u_{-i}| \right|_2^2$$

Thus $||diag(\alpha_i) U \Psi_\mu^{\inf} U^{-1}||_1 \leq 1 \Leftrightarrow (\alpha_1, \ldots, \alpha_n)$-spectral independence

- Scaling μ doesn't change interaction matrix, so only needs to show spectral independence

Overview

1. Motivation
   - Ising and hardcore model
   - Glauber dynamics
   - Multi-step down-up walks
   - Markov chain and mixing time
2. Entropic Independence
   - Definition
   - From fractional log-concavity to entropic independence
3. Tight mixing time for local walks
   - Local-to-global argument
   - Glauber dynamics for Ising/hardcore models

Glauber dynamics for hardcore models $(\lambda < \lambda_\Delta(1-\delta))$

- Hardcore model distribution is NOT fractionally log-concave
- Not spectrally independent when $\lambda_i > \lambda_\Delta$

Glauber dynamics for hardcore models $(\lambda < \lambda_\Delta(1 - \delta))$

- Hardcore model distribution is NOT fractionally log-concave
- Not spectrally independent when $\lambda_i > \lambda_\Delta$
- But is $O\left(\frac{1}{\delta}\right)$-spectral independence when $\lambda_i < \lambda_\Delta(1-\delta)\forall i$
- This implies a **restricted** form of entropic independence

$$\mathcal{D}_{KL}(\nu||\mu) \geq \alpha k \mathcal{D}_{KL}(\nu D_{k\to 1}||\mu D_{k\to 1})$$

for ν in a restricted class of distribution.

Glauber dynamics for hardcore models $(\lambda < \lambda_\Delta(1-\delta))$

Proof sketch:

1. Restricted entropic independence
2. (Restricted) entropy contraction for field-dynamics
   - Field dynamics: reduce sampling at $\lambda \equiv \lambda_\Delta$ to $\lambda \ll \lambda_\Delta$
   - Field dynamics can be viewed as multi-step down-up walk
   - Local-to-global arguments + restricted entropy contraction

Glauber dynamics for hardcore models $(\lambda < \lambda_\Delta(1-\delta))$

Proof sketch:

1. Restricted entropic independence
2. (Restricted) entropy contraction for field-dynamics
   - Field dynamics: reduce sampling at $\lambda \equiv \lambda_\Delta$ to $\lambda \ll \lambda_\Delta$
3. Restricted entropy contraction for Glauber dynamics
   - Each step of field dynamic is implementable by GD steps in easier regime $(\lambda \leq \frac{1}{\Delta})$
   - GD in easier regime has known entropy contraction.
   - Comparison between field-dynamics and Glauber dynamics

Subsequent works

- [Chen-Eldan'22,Chen-Feng-Yin-Zhang'22] use entropic independence to show full entropy contraction for hardcore model and other anti-ferromagnetic 2-spin systems.

Other applications